VIDEO ENCODING METHOD WITH SUPPORT FOR EDITING WHEN SCENE CHANGED

BACKGROUND OF THE INVENTION

Field of the invention

The invention relates to a video encoding method, especially to a video encoding method with support for editing when scene changed.

Description of the Related Art

In MPEG (Moving Pictures Experts Group), there is three picture types: I-picture, P-picture and B-picture. I-pictures are coded without reference to other pictures. They provide access points to the coded sequence where decoding can begin, but are coded with only moderate compression. P-pictures are coded more efficiently using motion compensated prediction from a past I-picture or P-picture and are generally used as a reference for further prediction. B-pictures provide the highest degree of compression but require both past and future reference pictures for motion compensation. B-pictures are never used as references for prediction. The organization of the three picture types in a sequence is very flexible. The choice is left to the encoder and will depend on the requirements of the application.

Because the B-pictures reference the past and future reference pictures, so the encoding process of the B-pictures has to be delayed until the future reference picture is coded. Therefore, the display order is different to the coding order. This is called the reordering of B-pictures.

In MPEG-1, there is a group-of-pictures (hereinafter called as GOP) structure used to enclose some pictures into a group for manipulation. A GOP contains one

I-picture, some P-pictures and some B-pictures. A GOP begins with an I-picture, and ends before the next I-pictures, in the coding order. In MPEG-2, the GOP structure becomes an option.

Generally, an encoder employs a fixed GOP structure. The size of a GOP is defined as N, and the distance between two reference pictures is defined as M. Fig. 1 illustrates a GOP with N = 15 and M = 3.

Typically, if the input signal for the encoder is in NTSC (National Television System Committee) format (29.97 fps), the GOP structure with N = 15 and M = 3 is used. If the input signal is in PAL (25 fps) or film format (24 fps), the GOP structure with N = 12 and M = 3 is used. These fixed default settings can achieve a good balance between the complexity of encoder and the coding performance for most videos.

Typically, the editing process would cut the whole video sequence into pieces based on the scene, and then rearrange them to form a new video sequence. If a video sequence is coded with a fixed pattern composed with only I- and P-pictures, like IPPPPIPPPP..., the situation is pretty simple. If a scene change occurred in an I-picture, the video sequence can be cut into two parts without any loss. If a scene change occurred in a P-picture, the former part is no problem, but the remaining part has to be re-encoder. The first P-picture has to be decoded and then re-encode to an I-picture. However, because the re-encoded I-picture differs from the original P-picture, there will be some error propagations. Re-encode the whole remaining part of the GOP until the next I-picture would be a better solution, but we would remind that re-encoding degrades the image quality significantly.

If there are B-pictures in the coded sequence, video editing becomes more

complex. Please reference to Fig. 2. If a scene change occurred in the picture just after the I-picture in the coding order, like the picture B_4, cutting from picture I_6 can separate the two scenes easily. However, even the picture P_3 and picture B_4 are belong to different scenes, there would be some macroblocks in picture B_4 and B_5 reference the picture P_3. Therefore the picture B_4 and B_5 have to be re-encoded with only referencing to the picture I_6. Discarding the pictures B_4 and B_5 is the easiest way, but losing the beginning some pictures of a scene would not be acceptable.

If a scene change occurred in the picture B_5, the former part and the remaining part of the GOP have some pictures to be re-encoded. The picture B_4 has to be re-encoded to a P-picture and then append to the former part. In the remaining part, the coded data of the picture B_4 is removed and the picture B_5 has to be re-encoded.

If a scene change occurred in the picture I_6, the remaining part has only to remove the coded data of the pictures B_4 and B_5. However, the former part requires a complicate process. One solution is to re-encode the picture B_5 to a P-picture, and then re-encode the picture B_4 by referencing to the pictures P_3 and B_5. Another solution is to change the two B-pictures to two P-pictures.

If a scene change occurred in the picture B_7, the former part needs no any process, but a new I-picture has to be generated for the remaining part. A choice is to change the picture B_7 to an I-picture, and then re-encode the remaining GOP. However, because the B-pictures usually coded with a lower quality than the I- and P-pictures, a better choice would be to change the picture P_9 to an I-picture, and re-encode the remaining GOP. The pictures B_5 and B_6 are B-pictures with only backward reference. This method can also reduce the number of P-pictures to

reduce the error caused by referencing to a re-encoded picture.

If a scene change occurred in the picture B_8, the former part has only to re-encode the picture B_7 to a P-picture. The remaining part can change the picture P_9 to an I-picture and then re-encode the remaining GOP.

Finally, if a scene change occurred in the picture P_9, the former part is processed like the situation of picture I_6. For the remaining part, the picture P_9 has to be changed to an I-picture, and then re-encode the remaining GOP.

Therefore, all the other situations can be processed like the methods described above, even if the number of B-pictures between two reference pictures increases to three or more.

Generally, the I-pictures are designed for the purpose of random access and preventing of error propagation. The P-pictures use the motion compensation to remove the temporal redundancy between the current picture and the reference picture to improve the compression performance. However, if there is almost no temporal redundancy between the current picture and the reference picture, for example a scene change, coding a picture as a P-picture can't obtain any benefit. In this case, an I-picture can achieve the same coding quality with fewer bits. Therefore, an encoder has to detect the existence of a scene change and then start a new GOP. There is already many researches focus on the scene change detection and then how to adjust the rate control algorithm. A general idea is to detect the difference of the current picture and the reference picture from the result of motion estimation. If more than a percentage of macroblocks select the intra-coded mode, the encoder can decide that there is only few temporal redundancy existed, and therefore a scene change can be detected.

However, if the encoder just start a new GOP when detect a scene change but with no other effort, the re-encoding of some pictures would be unavoidable when the video sequence is being editing, as we described above.

SUMMARY OF THE INVENTION

In view of the above-mentioned problems, an object of the invention is to provide a video encoding method with support for editing when scene changed.

To achieve the above-mentioned object, the video encoding method with support for editing when scene changed of the present invention encodes the pictures by the coding order when there are not scenes changed and encodes the pictures by a special coding process when there are scenes changed. Because the video encoding method encodes the pictures with considering the states of scenes changed and generates a new GOP when a scene change occurred, the video sequence can be cut into two parts by an image editing process without re-encoding.

Therefore, the video can be edited without any loss and the editing performance of the editing process can be better.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates a GOP with N = 15 and M = 3.

Fig. 2 illustrates a video sequence with a B-picture.

Fig. 3 illustrates a GOP with a fixed structure.

Fig. 4 illustrates a first example of GOP with display order and coding order.

Fig. 5 illustrates a second example of GOP with display order and coding order.

Fig. 6 illustrates a third example of GOP with display order and coding order.

Fig. 7 shows the flowchart of the video encoding method with support for

editing when scene changed of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The video encoding method with support for editing when scene changed of the invention will be described with reference to the accompanying drawings.

First, an encoder using the video encoding method with support for editing when scene changed of the present invention also has to include the scene change detecting function. The scene change detecting function has to be applied in the display order. This is because the encoder has to know where scene changes and then encodes the pictures before and after the scene change into two GOP.

Before a scene change is detected, the encoder encodes the video sequence with a fixed GOP structure. Once a scene change is detected, the encoder decides how to encode the following pictures based on the type and position in a GOP of the just coded pictures. Please note that because the B-pictures have to be coded just after the future reference picture being coded, a scene change can be detected far before the coding actually happen. Fig. 2 depicts an example. The pictures are captured and stored into a buffer by the display order. The picture B_4 and B_5 are captured but can't be encode until the picture I_6 is coded. Assume that the encoder can encode a picture in each period of capturing a picture. The picture I_6 is captured and then encoded in the same period. In the next period, the picture B_4 is encoded while the picture B_7 is being captured. The picture B_5 is encoded in the same period that the picture B_7 is captured. The picture P_9 needs only the picture I_6 to reference, so it can be captured and encoded in the same period.

An encoder encodes the video sequence with a fixed GOP structure that the distance between two reference pictures is defined as M and a reference picture (I- or

P-picture) is represented as an R. The first B-picture (in the display order) after the forward reference picture R^X is called B^X_1, the second B-picture is called B^X_2, and so on. The final one before the backward reference picture is called B^X_{M-1}. Fig. 3 illustrates an example of the GOP in the display order and coding order.

The process methods for the scene change occurred at different location are described as following.

A. A scene change occurred in the first B-picture

If there is no scene change occurred in the pictures from B^A_1 to R^B, the picture $B^A_1 \sim B^A_{M-1}$ is captured and stored until the picture R^B is captured and coded. If the scene changed in the picture B^B_1, the pictures until R^B would belong to the former GOP and the pictures from B^B_1 would belong to a new GOP. After coding the picture B^A_{M-1}, if the encoder starts a new GOP and encodes the following pictures without referencing to the picture R^B, it can completely separate the video sequence into two parts. An editing process can cut the video sequence from the new GOP without any re-encoding.

There are two strategies to start a new GOP. One is to start a fixed GOP structure from I-picture. In the above example, the original picture B^B_1 is changed to an I-picture R^C, the following M-1 pictures are B-picture $B^C_1 \sim B^C_{M-1}$, the next is a P-picture R^D, then are the M-1 B-pictures, and so on. Fig. 4 illustrates an example of this case.

However, a new GOP need not be started with an I-picture in the display order. By observing the coding order in Fig. 4, we can find that there is no B-pictures between the picture R^C and R^D. B-pictures can be coded with lower quality and save the bit rate for the I-pictures and P-pictures. If there are too many reference

pictures in a short duration, the result is that each reference picture can't obtain enough bits to achieve a higher quality. Therefore, the second strategy of starting a new GOP is trying to maintain the ratio on the number of B-pictures and reference pictures. The first M-1 pictures of the new GOP are B-pictures, the next is an I-pictures, following by other M-1 B-pictures, and then a P-picture, and so on. Fig. 5 illustrates an example of this case.

Seems that the picture type of each picture is remaining the same as no scene change occurred. Actually, the difference is that the picture $B^B_1 \sim B^B_{M-1}$ have only backward reference to the picture R^C. In fact, there may not be M-1 B-pictures before the picture R^C, and can be adjust freely.

B. A scene change occurred in the second B-picture

Please reference to Fig. 3. If a scene change occurred in the picture B^B_2, the picture B^B_1 belongs to the former GOP and the pictures from B^B_2 form a new GOP. The new GOP can be encoded with the same method described in subsection A.

A GOP can be ended by a reference picture. Therefore the picture B^B_1 must be encoded as a reference picture. There is no reason to not encode the picture B^B_1 as a P-picture but an I-picture. Fig. 6 illustrates an example of this case.

C. A scene change occurred in the n-th B-picture

If a scene change occurred in the n-th B-picture after the reference picture R^X, $2 \leq n \leq M-1$, the pictures until B^X_{n-1} belong to the former GOP and the pictures from B^X_n form a new GOP. The new GOP is encoded with the same method described in subsection A.

Based on the method described in section B, the encoder will encode the picture B^X_{n-1} as a P-picture, and the pictures $B^X_1 \sim B^X_{n-2}$ (if any) are encoded as

B-pictures by referencing to the picture R^X and the new generated P-picture.

D. A scene change occurred in a reference picture

Please reference to Fig. 3. If a scene change occurred in the picture R^B, the picture B^A_{M-1} belongs to the former GOP and the pictures from R^B form a new GOP. The former GOP can be coded by the method described in subsection C. The new GOP could be encoded with the same method described in subsection A.

Finally, we can simplify an algorithm for an encoder to encode the pictures before and after a scene change separately to make the video editing of scenes without any re-encoding. Fig. 7 shows the flowchart of the video encoding method with support for editing when scene changed of the present invention.

Step 702: Capture the picture PIC_n in the display order and detect the scene change.

Step 704: If there is no scene change in the picture PIC_n, the flowchart jumps to step S706. If there is a scene change in the picture PIC_n, the flowchart jumps to step S708.

Step 706: Code pictures in the coding order and jump back to step S702.

Step 708: If the picture PIC_{n-1} is not coded as a reference picture, jump to step S710. If the picture PIC_{n-1} is coded as a reference picture, jump to step S716.

Step 710: If there are B-pictures preceding a previous reference picture, finish coding the B-pictures.

Step S712: Encode the picture PIC_{n-1} as a P-picture.

Step S714: If there are B-pictures preceding the picture PIC_{n-1}, coding the B-pictures and jump to step S718.

Step S716: If there are B-pictures preceding the picture PIC_{n-1}, coding the